UpSetR motif Enh DAR DMRs intersection within each Layer: show how similar they are

-We are checking if Enh DAR DMR motifs are lineage-specific within laeyr, e.g. contain mostly different between each other motifs and most similar within any Layer

library(tidyverse)

## -- Attaching packages ---------------------------------- tidyverse 1.3.0 --

## v ggplot2 3.3.0 v purrr 0.3.3  
## v tibble 3.0.0 v dplyr 0.8.5  
## v tidyr 1.0.2 v stringr 1.4.0  
## v readr 1.3.1 v forcats 0.5.0

## -- Conflicts ------------------------------------- tidyverse\_conflicts() --  
## x dplyr::filter() masks stats::filter()  
## x dplyr::lag() masks stats::lag()

library(ggplot2)  
library(gridExtra)

##   
## Attaching package: 'gridExtra'

## The following object is masked from 'package:dplyr':  
##   
## combine

library(tidyr)  
library(ggpubr)

## Loading required package: magrittr

##   
## Attaching package: 'magrittr'

## The following object is masked from 'package:purrr':  
##   
## set\_names

## The following object is masked from 'package:tidyr':  
##   
## extract

library("readxl")  
library(datasets)   
library(UpSetR)  
library(VennDiagram)

## Loading required package: grid

## Loading required package: futile.logger

##   
## Attaching package: 'VennDiagram'

## The following object is masked from 'package:ggpubr':  
##   
## rotate

library(repurrrsive) #for list length

-read motifs Excel files: Enh DMR DAR

-read motifs Enhs

#-----------------Enhancers weis  
  
#read\_excel("C:/Users/ia1/Documents/irina\_docs/EpiGeVar/MayJune2020\_egv/DARs\_Carine/wei\_ecto.xlsx")->motifs\_EnhancerEc1  
#motifs\_EnhancerEc1  
  
read\_excel("data\_motifs\_Carine/wei\_ecto.xlsx")->motifs\_EnhancerEc1  
motifs\_EnhancerEc1

## # A tibble: 44 x 8  
## TF P logP `q-Ben` `#Tar\_Motif` percTar\_Motif `#bg\_Motif`  
## <chr> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>  
## 1 OCT4~ 1.00e-25 -59.3 0 62 0.214 2018.  
## 2 Sox2~ 1.00e-13 -31.7 0 116 0.4 9925   
## 3 Sox4~ 1.00e-12 -28 0 109 0.376 9511.  
## 4 Sox2~ 1.00e-12 -28 0 181 0.624 20277.  
## 5 Sox3~ 1.00e-10 -25.2 0 167 0.576 18543.  
## 6 Sox1~ 1.00e- 9 -21.6 0 87 0.3 7577.  
## 7 Sox6~ 1.00e- 9 -20.9 0 150 0.517 16732.  
## 8 Oct4~ 1.00e- 8 -20.4 0 62 0.214 4663.  
## 9 Sox1~ 1.00e- 8 -20 0 154 0.531 17559   
## 10 Sox7~ 1.00e- 8 -19 0 47 0.162 3134   
## # ... with 34 more rows, and 1 more variable: perc\_bg\_Motif <dbl>

read\_excel("data\_motifs\_Carine/wei\_endo.xlsx")->motifs\_EnhancerEn1  
#motifs\_EnhancerMes1  
  
read\_excel("data\_motifs\_Carine/wei\_meso.xlsx")->motifs\_EnhancerMes1  
#motifs\_EnhancerMes1

-read motifs DMR

#-----------------DMRs   
  
read\_excel("data\_motifs\_Carine/wei\_ecto.xlsx")->motifs\_DMREc1  
#motifs\_DMRMes1  
  
read\_excel("data\_motifs\_Carine/wei\_endo.xlsx")->motifs\_DMREn1  
#motifs\_DMRMes1  
  
read\_excel("data\_motifs\_Carine/wei\_meso.xlsx")->motifs\_DMRMes1  
#motifs\_DMRMes1

-read motifs DARs

read\_excel("data\_motifs\_Carine/DAR\_ecto.xlsx")->motifs\_DAREc1  
  
read\_excel("data\_motifs\_Carine/DAR\_endo.xlsx")->motifs\_DAREn1  
#motifs\_DARMes  
  
  
read\_excel("data\_motifs\_Carine/DAR\_meso.xlsx")->motifs\_DARMes1  
#motifs\_DARMes1

-B. set quantile thresholds for each Layer regions

-B.DMR quantiles DMR

quantile(motifs\_DMREc1$P)->q\_dmr\_ec  
quantile(motifs\_DMREn1$P)->q\_dmr\_en  
quantile(motifs\_DMRMes1$P)->q\_dmr\_m  
q\_dmr\_ec

## 0% 25% 50% 75% 100%   
## 1.00e-25 7.75e-08 1.00e-04 1.00e-02 1.00e-02

q\_dmr\_en

## 0% 25% 50% 75% 100%   
## 1e-51 1e-13 1e-03 1e-02 1e-02

q\_dmr\_m

## 0% 25% 50% 75% 100%   
## 1e-17 1e-07 1e-05 1e-03 1e-02

-set thr DMR, quantiles Q75%

thrPe\_ec2<-q\_dmr\_ec[4]  
thrPe\_en2<-q\_dmr\_en[3]# to make a little smaller  
thrPe\_m2<-q\_dmr\_m[4]

-make individual lists for DMRs: get the functions

source("make\_filterP\_list\_dmrEc.R")  
source("make\_filterP\_list\_dmrEn.R")  
source("make\_filterP\_list\_dmrM.R")

* upsetR plot DMRs -how similar dmr motifs are bw layers, thrP quantile

#-----------------------Q75 thr  
  
vdmrec2<-make\_filterP\_list\_dmrEc(motifs\_DMREc1,thrPe\_ec2)  
#library(repurrrsive)  
vdmrec2%>%  
map(length)

## $motifs\_DMREc  
## [1] 44

#typeof(vdmrec)  
  
vdmren2<-make\_filterP\_list\_dmrEn(motifs\_DMREn1,thrPe\_en2)  
vdmren2%>%  
map(length)

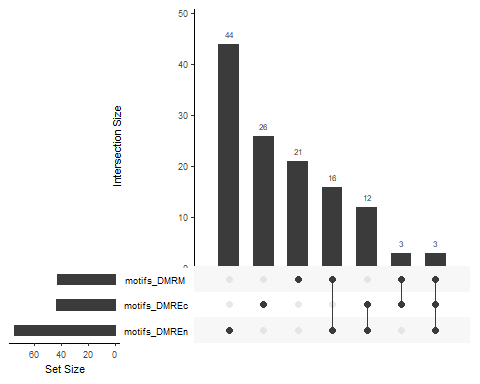
## $motifs\_DMREn  
## [1] 75

vdmrm2<-make\_filterP\_list\_dmrM(motifs\_DMRMes1,thrPe\_m2)  
vdmrm2%>%  
map(length)

## $motifs\_DMRM  
## [1] 43

* upsetR plot DMRs -how similar dmr motifs are bw layers, thrP quantile

listik\_dmr\_EcEnM2=c(vdmrec2,vdmren2,vdmrm2)  
upset(fromList(listik\_dmr\_EcEnM2),empty.intersections = "on")



#----------list of motifs  
#vdmrec2  
#vdmren2

-B.DAR quantiles DAR

quantile(motifs\_DAREc1$P)->q\_dar\_ec  
quantile(motifs\_DAREn1$P)->q\_dar\_en  
quantile(motifs\_DARMes1$P)->q\_dar\_m  
q\_dar\_ec

## 0% 25% 50% 75% 100%   
## 1e-90 1e-10 1e-05 1e-03 1e-02

q\_dar\_en

## 0% 25% 50% 75% 100%   
## 1e-84 1e-17 1e-06 1e-03 1e-02

q\_dar\_m

## 0% 25% 50% 75% 100%   
## 1e-49 1e-13 1e-05 1e-03 1e-02

-set thr DAR, quantiles Q25%

thrPe\_ec2<-q\_dar\_ec[2]  
thrPe\_en2<-q\_dar\_en[2]# to make a little smaller  
thrPe\_m2<-q\_dar\_m[2]

-make individual lists for DARs: get the functions

source("make\_filterP\_list\_darEc.R")  
source("make\_filterP\_list\_darEn.R")  
source("make\_filterP\_list\_darM.R")

* upsetR plot DARs -how similar dmr motifs are bw layers, thrP quantile

#-----------------------Q75 thr  
  
vdarec2<-make\_filterP\_list\_darEc(motifs\_DAREc1,thrPe\_ec2)  
#library(repurrrsive)  
vdarec2%>%  
map(length)

## $motifs\_DAREc  
## [1] 46

#typeof(vdarec)  
  
vdaren2<-make\_filterP\_list\_darEn(motifs\_DAREn1,thrPe\_en2)  
vdaren2%>%  
map(length)

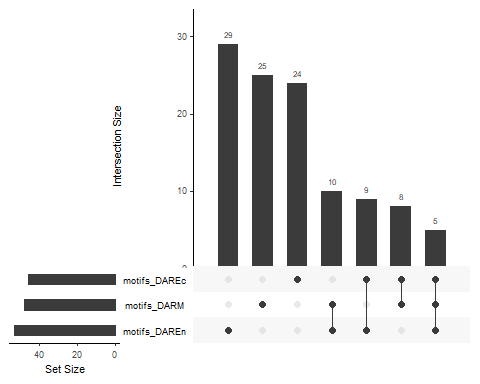
## $motifs\_DAREn  
## [1] 53

vdarm2<-make\_filterP\_list\_darM(motifs\_DARMes1,thrPe\_m2)  
vdarm2%>%  
map(length)

## $motifs\_DARM  
## [1] 48

* upsetR plot DARs -how similar dar motifs are bw layers, thrP quantile

#---------------------should be different mostly  
listik\_dar\_EcEnM2=c(vdarec2,vdaren2,vdarm2)  
upset(fromList(listik\_dar\_EcEnM2),empty.intersections = "on")

 -B.Wei Enhancers

-quantiles Wei

quantile(motifs\_EnhancerEc1$P)->q\_w\_ec  
quantile(motifs\_EnhancerEn1$P)->q\_w\_en  
quantile(motifs\_EnhancerMes1$P)->q\_w\_m  
q\_w\_ec

## 0% 25% 50% 75% 100%   
## 1.00e-25 7.75e-08 1.00e-04 1.00e-02 1.00e-02

q\_w\_en

## 0% 25% 50% 75% 100%   
## 1e-51 1e-13 1e-03 1e-02 1e-02

q\_w\_m

## 0% 25% 50% 75% 100%   
## 1e-17 1e-07 1e-05 1e-03 1e-02

* change thr, quantiles Q75% Wei

thrPe\_ec2<-q\_w\_ec[4]  
thrPe\_en2<-q\_w\_en[3]# to make a little smaller  
thrPe\_m2<-q\_w\_m[4]

-make ind lists for Enhancers Wei

source("make\_filterP\_list\_enhEc.R")  
source("make\_filterP\_list\_enhEn.R")  
source("make\_filterP\_list\_enhM.R")

-quntile P-threshold Wei

#-----------------------Q75 thr Wei  
  
vwec2<-make\_filterP\_list\_enhEc(motifs\_EnhancerEc1,thrPe\_ec2)  
#library(repurrrsive)  
vwec2%>%  
map(length)

## $motifs\_EnhancerEc  
## [1] 44

#typeof(vwec)  
  
vwen2<-make\_filterP\_list\_enhEn(motifs\_EnhancerEn1,thrPe\_en2)  
vwen2%>%  
map(length)

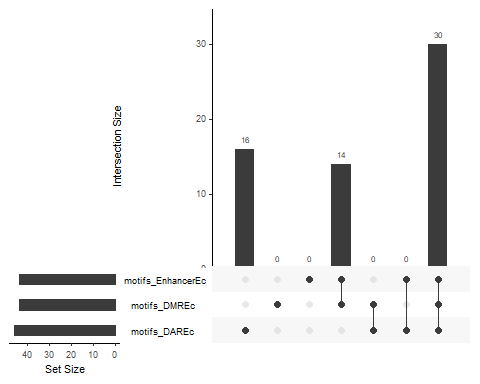
## $motifs\_EnhancerEn  
## [1] 75

vwm2<-make\_filterP\_list\_enhM(motifs\_EnhancerMes1,thrPe\_m2)  
vwm2%>%  
map(length)

## $motifs\_EnhancerM  
## [1] 43

-Ectoderm: Enh DAR DMR regions

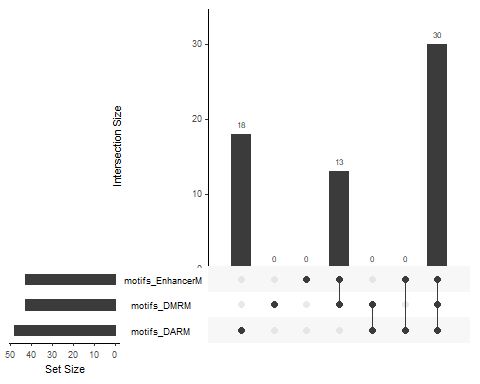
listik\_w\_dar\_dmr\_Ec2=c(vwec2,vdarec2,vdmrec2)  
upset(fromList(listik\_w\_dar\_dmr\_Ec2),empty.intersections = "on")

 -get elements Ectoderm

overlapEct<-calculate.overlap(listik\_w\_dar\_dmr\_Ec2)  
#overlapEct

-Mesoderm: intersection of Enh DAR DMR

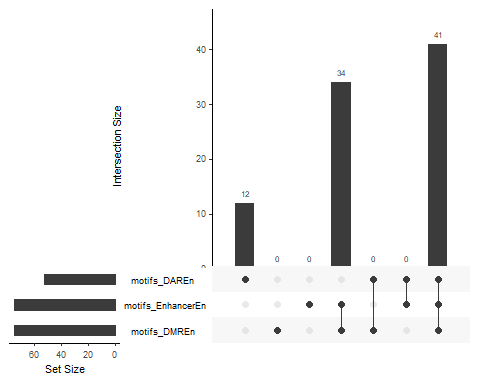
listik\_w\_dar\_dmr\_M2=c(vwm2,vdarm2,vdmrm2)  
upset(fromList(listik\_w\_dar\_dmr\_M2),empty.intersections = "on")

 -get upset elements Mesoderm: inresect, distinct, paiwise

overlapMes<-calculate.overlap(listik\_w\_dar\_dmr\_M2)  
#overlapMes

-Endoderm within: intersection of Enh DAR DMR

listik\_w\_dar\_dmr\_En2=c(vwen2,vdaren2,vdmren2)  
upset(fromList(listik\_w\_dar\_dmr\_En2),empty.intersections = "on")

 -get elements Endoderm

overlapEnd<-calculate.overlap(listik\_w\_dar\_dmr\_En2)  
#overlapEnd